NCBI

Nucleotide

PubMed	Nucleotide	Protein	Genome	Structure	PopSet	Taxonomy	OMIM	Books

Search Nucleotide ▼ for _____ | Go | Clear

Limits   Preview/Index   History   Clipboard   Details

| Display | default | ▼ | Save | Text | Add to Clipboard |

☐ **1: K01760. HSV1 (KOS) glycop...[gi:330082]**   Related Sequences, Protein, PubMed, Taxonomy

```
LOCUS       HS1GB                   3758 bp    DNA     linear   VRL 02-AUG-1993
DEFINITION  HSV1 (KOS) glycoprotein B gene, complete cds.
ACCESSION   K01760
VERSION     K01760.1  GI:330082
KEYWORDS    glycoprotein; glycoprotein B.
SOURCE      Herpes simplex virus type 1 (strain KOS) DNA, maintained on HEL
            cells.
  ORGANISM  human herpesvirus 1
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae; Simplexvirus.
REFERENCE   1  (bases 1 to 3758)
  AUTHORS   Bzik,D.J., Fox,B.A., DeLuca,N.A. and Person,S.
  TITLE     Nucleotide sequence specifying the glycoprotein gene, gB, of Herpes
            simplex virus type 1
  JOURNAL   Virology 133, 301-314 (1984)
  MEDLINE   84174058
REFERENCE   2  (bases 1 to 3321)
  AUTHORS   Bzik,D.J., Debroy,C., Fox,B.A., Pederson,N.E. and Person,S.
  TITLE     The nucleotide sequence of the gB glycoprotein gene of HSV-2 and
            comparison with the corresponding gene of HSV-1
  JOURNAL   Virology 155, 322-333 (1986)
  MEDLINE   87071654
REFERENCE   3  (bases 1 to 3311)
  AUTHORS   Pederson,N.
  JOURNAL   Unpublished (1987)
COMMENT     [2]  revises [1].
            [3]  revises [1],[2].
            Draft entry and computer-readable sequence for [3] kindly provided
            by N.Pederson  14-APR-1987.
            There are 4 antigenically distinct glycoproteins in HSV1 infected
            cells -- gB, gC, gD and gE. The glycoprotein formerly known as gA
            is now know to be a precursor to gB and has been renamed pgB. gB is
            the only glycoprotein that is known to be required for viral
            growth. The gB mRNA probably runs about from bases 257 to 3300.
            Strand shown runs in direction opposite to that of the standard HSV
            map, from mu .370 to .345.
FEATURES             Location/Qualifiers
     source          1..3758
                     /organism="human herpesvirus 1"
                     /db_xref="taxon:10298"
     CDS             548..3262
                     /note="glycoprotein B precursor (pgB)"
                     /codon_start=1
                     /protein_id="AAA45774.1"
                     /db_xref="GI:330083"
                     /translation="MHQGAPSWGRRWFVVWALLGLTLGVLVASAAPTSPGTPGVAAAT
                     QAANGGPATPAPPPLGAAPTGDPKPKKNKKPKNPTPPRPAGDNATVAAGHATLREHLR
                     DIKAENTDANFYVCPPPTGATVVQFEQPRRCPTRPEGQNYTEGIAVVFKENIAPYKFK
                     ATMYYKDVTVSQVWFGHRYSQFMGIFEDRAPVPFEEVIDKINAKGVCRSTAKYVRNNL
                     ETTAFHRDDHETDMELKPANAATRTSRGWHTTDLKYNPSRVEAFHRYGTTVNCIVEEV
```

[handwritten annotations: "HSV-1 gB1", "HSV2gB"]

```
                    DARSVYPYDEFVLATGDFVYMSPFYGYREGSHTEHTTYAADRFKQVDGFYARDLTTKA
                    RATAPTTRNLLTTPKFTVAWDWVPKRPSVCTMTKWQEVDEMLRSEYGGSFRFSSDAIS
                    TTFTTNLTEYPLSRVDLGDCIGKDARDAMDRIFARRYNATHIKVGQPQYYQANGGFLI
                    AYQPLLSNTLAELYVREHLREQSRKPPNPTPPPPGASANASVERIKTTSSIEFARLQF
                    TYNHIQRHVNDMLGRVAIAWCELQNHELTLWNEARKLNPNAIASVTVGRRVSARMLGD
                    VMAVSTCVPVAADNVIVQNSMRISSRPGACYSRPLVSFRYEDQGPLVEGQLGENNELR
                    LTRDAIEPCTVGHRRYFTFGGGYVYFEEYAYSHQLSRADITTVSTFIDLNITMLEDHE
                    FVPLEVYTRHEIKDSGLLDYTEVQRRNQLHDLRFADIDTVIHADANAAMFAGLGAFFE
                    GMGDLGRAVGKVVMGIVGGVVSAVSGVSSFMSNPFGALAVGLLVLAGLAAAFFAFRYV
                    MRLQSNPMKALYPLTTKELKNPTNPDASGEGEEGGDFDEAKLAEAREMIRYMALVSAM
                    ERTEHKAKKKGTSALLSAKVTDMVMRKRRNTNYTQVPNKDGDADEDDL"
     variation      1364
                    /note="g in KOS; a in mutant tsJ20"
     variation      1676
                    /note="c in KOS; t in mutant tsJ12"
BASE COUNT      717 a    1286 c    1128 g     627 t
ORIGIN       295 bp upstream of XhoI site; left half of BamHI-G frag at .370 mu.
        1 gtcaacgggc ccctctttga tcactccacc cacagcttcg cccagccccc caacaccgcg
       61 ctgtattaca gcgtcgagaa cgtggggctc ctgccgcacc tgaaggagga gctcgcccgg
      121 ttcatcatgg gggcgggggg ctcgggtgct gattgggccg tcagcgaatt tcagaggttt
      181 tactgttttg acggcatttc cggaataacg cccactcagc gcgccgcctg gcgatatatt
      241 cgcgagctga ttatcgccac cacactcttt gcctcggtct accggtgcgg ggagctcgag
      301 ttgcgccgcc cggactgcag ccgcccgacc tccgaaggtc gttaccgtta cccgcccggc
      361 gtatatctca cgtacgactc cgactgtccg ctggtggcca tcgtcgagag cgcccccgac
      421 ggctgtatcg gcccccggtc ggtcgtggtc tacgaccgcg acgtttttctc gatcctctac
      481 tcggtcctcc agcacctcgc ccccaggcta cctgacgggg ggcacgacgg gcccccgtag
      541 tcccgccatg caccagggcg cccCcctcgtg ggggcgccgg tggttcgtcg tatgggcgct
      601 cttggggttg acgctggggg tcctggtggc gtcggcggct ccgacttccc ccggcacgcc
      661 tggggtcgcg gccgcgaccc aggcggcgaa cgggggccct gccactccgg cgccgccgcc
      721 ccttggcgcc gccccaacgg gggacccgaa accgaagaag aacaaaaaac cgaaaaaccc
      781 aacgccacca cgccccgccg gcgacaacgc gaccgtcgcc gcgggccacg ccaccctgcg
      841 cgagcacctg cgggacatca aggcggagaa caccgatgca aactttttacg tgtgcccacc
      901 ccccacgggc gccacggtgg tgcagttcga gcagccgcgc cgctgcccga cccggcccga
      961 gggtcagaac tacacggagg gcatcgcggt ggtcttcaag gagaacatcg ccccgtacaa
     1021 gttcaaggcc accatgtact acaaagacgt caccgtttcg caggtgtggt tcggccaccg
     1081 ctactcccag tttatgggga tctttgagga ccgcgccccc gtccccttcg aggaggtgat
     1141 cgacaagatc aacgccaagg gggtctgtcg gtccacggcc aagtacgtgc gcaacaacct
     1201 ggagaccacc gcgtttcacc gggacgacca cgagaccgac atggagctga aaccggccaa
     1261 cgccgcgacc cgcacgagcc ggggctggca caccaccgac ctcaagtaca acccctcgcg
     1321 ggtggaggcg ttccaccggt acgggacgac ggtaaactgc atcgtcgagg aggtggacgc
     1381 gcgctcggtg tacccgtacg acgagtttgt gctggcgact ggcgactttg tgtacatgtc
     1441 cccgtttttac ggctaccggg aggggtcgca caccgaacac accacgtacg ccgccgaccg
     1501 cttcaagcag gtcgacggct tctacgcgcg cgacctcacc accaaggccc gggccacggc
     1561 gccgaccacc cggaacctgc tcacgacccc caagttcacc gtggcctggg actgggtgcc
     1621 aaagcgcccg tcggtctgca ccatgaccaa gtggcaggaa gtggacgaga tgctgcgctc
     1681 cgagtacggc ggctccttcc gattctcctc cgacgccata tccaccacct tcaccaccaa
     1741 cctgaccgag tacccgctct cgcgcgtgga cctggggggac tgcatcggca aggacgcccg
     1801 cgacgccatg gaccgcatct tcgcccgcag gtacaacgcg acgcacatca aggtgggcca
     1861 gccgcagtac taccaggcca atggggggctt tctgatcgcg taccagcccc ttctcagcaa
     1921 cacgctcgcg gagctgtacg tgcgggaaca cctccgagag cagagccgca agccccaaa
     1981 ccccacgccc ccgccgcccg gggccagcgc caacgcgtcc gtggagcgca tcaagaccac
     2041 ctcctccatc gagttcgccc ggctgcagtt tacgtacaac cacatacagc gccatgtcaa
     2101 cgatatgttg ggccgcgttg ccatcgcgtg gtgcgagcta cagaatcacg agctgaccct
     2161 gtggaacgag gccccgcaag ctgaacccca cgccatcgcc tcggtcaccg tgggccggcg
     2221 ggtgagccgcg cggatgctcg gcgacgtgat ggccgtctcc acgtgcgtgc cggtcgccgc
     2281 ggacaacgtg atcgtccaaa actcgatgcg catcagctcg cggcccgggg cctgctacag
     2341 ccgcccCctg gtcagctttc ggtacgaaga ccagggcccg ttggtcgagg ggcagctggg
     2401 ggagaacaac gagctgcggc tgacgcgcga tgcgatcgag ccgtgcaccg tgggacaccg
     2461 gcgctacttc accttcggtg ggggctacgt gtacttcgag gagtacgcgt actcccacca
     2521 gctgagccgc gccgacatca ccaccgtcag caccttcatc gacctcaaca tcaccatgct
     2581 ggaggatcac gagtttgtcc ccctggaggt gtacacccgc cacgagatca aggacagcgg
     2641 cctgctggac tacacggagg tccagcgccg caaccagctg cacgacctgc gcttcgccga
     2701 catcgacacg gtcatccacg ccgacgccaa cgccgccatg ttcgcgggcc tgggcgcgtt
     2761 cttcgagggg atgggcgacc tggggcgcgc ggtcggcaag gtggtgatgg gcatcgtggg
     2821 cggcgtggta tcggccgtgt cgggcgtgtc ctccttcatg tccaaccccct ttggggcgct
```

```
2881 ggccgtgggt ctgttggtcc tggccggcct ggcggcggcc ttcttcgcct ttcgttacgt
2941 catgcggctg cagagcaacc ccatgaaggc cctgtaccct ctaaccacca aggagctcaa
3001 gaaccccacc aacccggacg cgtccgggga gggcgaggag ggcggcgact ttgacgaggc
3061 caagctagcc gaggccaggg agatgatacg gtacatggcc ctggtgtcgg ccatggagcg
3121 cacggaacac aaggccaaga agaagggcac gagcgcgctg ctcagcgcca aggtcaccga
3181 catggtcatg cgcaagcgcc gcaacaccaa ctacacccaa gttcccaaca aagacggtga
3241 cgccgacgag gacgacctgt gacgggggggt ttgttgtaaa taaaaaccac gggtgttaaa
3301 ccgcatgcgc atcttttggt ttttttgttt ggtcagcctt ttgtgtgtgt gtgggaagaa
3361 agaaaaagga acacataaac tcccccgggt gtccgcggcc tgtttcctct ttcctttccc
3421 gtgacaaaac ggaccccctt ggtcagtgcc gatttcctcc cccccacgcc ttcctccacg
3481 tcaaaggctt ttgcattgta aagctacccg cctacccgcc cctcccaata aaaaaaaaag
3541 aacatacacc aatgggtctt atttggtatt acctggttta tttaaaaaga tatacagtaa
3601 gacatcccat ggtaccaaag accggggcga atcagcgggc ccccatcatc tgagagacga
3661 acaaatcggc ggcgcgggcc gtgtcaacgt ccacgtgtgc tgcgctgctg gcgttgacaa
3721 ggccccggcc tccgcgttgg atgctccggt tgggatcc
//
```

Revised: October 24, 2001.

NCBI

PubMed

National Library of Medicine NLM

| PubMed | Nucleotide | Protein | Genome | Structure | PopSet | Taxonomy | OMIM | Books |

Search PubMed for [Go] [Clear]

Limits   Preview/Index   History   Clipboard   Details

Display   Citation   Sort   Save   Text   Clip Add   Order

Entrez
PubMed

☐ 1: Virology 1984 Mar;133(2):301-14

Related Articles, Nucleotide, Protein, NEW **Book**
LinkOu

PubMed
Services

Nucleotide sequence specifying the glycoprotein gene, gB, of herpes simplex virus type 1.

Bzik DJ, Fox BA, DeLuca NA, Person S.

The nucleotide sequence thought to specify the glycoprotein gene, gB, of the KOS strain of herpes simplex virus type 1 (HSV-1) has been determined. A 3.1-kilobase (kb), viral-specified RNA was mapped to the left half of the BamHI-G fragment (0.345 to 0.399 map units). TATA, CAT-box, and possible mRNA start sequences characteristic of HSV-1 genes are found near 0.368 map units. The first available ATG codon is at 0.366 and the first in-phase chain terminator at 0.348 map units. A polyA-addition signal (AATAAA) occurs 17 nucleotides past the chain terminator. Translation of these sequences would yield a 100.3-kilodalton (kDa) polypeptide characterized by a 5' signal sequence, nine N-linked saccharide addition sites, a strongly hydrophobic membrane-spanning sequence, and a highly charged 3' cytoplasmic anchor sequence. Two mutants of KOS, tsJ12 and tsJ20, that are temperature-sensitive for viral growth and for the production of gB, have been physically mapped to 0.357 to 0.360 and 0.360 to 0.364 map units, respectively (DeLuca et al., in preparation). The nucleotide sequence of the mutants was determined in these regions. In both cases a single amino acid replacement within the 100.3-kDa polypeptide is predicted from the sequence analysis.

Related
Resources

MeSH Terms:
- Amino Acid Sequence
- Base Sequence
- Cell Line
- DNA Restriction Enzymes
- DNA, Recombinant/analysis
- Genes, Structural*
- Genes, Viral*
- Human
- Lung/embryology
- Nucleic Acid Hybridization
- Simplexvirus/genetics*
- Support, Non-U.S. Gov't
- Support, U.S. Gov't, P.H.S.

Substances:
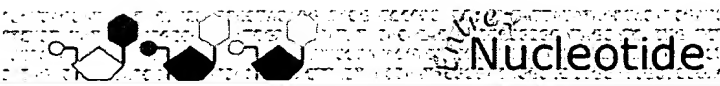- DNA Restriction Enzymes
- DNA, Recombinant

Secondary source id:
- GENBANK/K01760

PMID: 6324454 [PubMed - indexed for MEDLINE]

Display   Citation   ▽   Sort   ▽ **Save**   Text   Clip Add   Order

℅ NCBI

Nucleotide

PubMed	Nucleotide	Protein	Genome	Structure	PopSet	Taxonomy	OMIM	Books

Search | Nucleotide | ▼ | for | _____ | Go | Clear

Limits   Preview/Index   History   Clipboard   Details

Display | default | ▼ | Save | Text | Add to Clipboard

☐ **1: M21629. Herpes simplex vi...[gi:330109]**   Related Sequences, Protein, PubMed, Taxonomy

```
LOCUS       HS1GLYB                 9756 bp    DNA     linear   VRL 01-JUL-1994
DEFINITION  Herpes simplex virus type 1 glycoprotein B gene (gB-1), complete
            cds.
ACCESSION   M21629
VERSION     M21629.1  GI:330109
KEYWORDS    glycoprotein B.
SOURCE      Herpes simplex virus type 1 DNA.
  ORGANISM  human herpesvirus 1
            Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
            Alphaherpesvirinae; Simplexvirus.
REFERENCE   1  (bases 1 to 9756)
  AUTHORS   Hammerschmidt,W., Conraths,F., Mankertz,J., Pauli,G., Ludwig,H. and
            Buhk,H.J.
  TITLE     Conservation of a gene cluster including glycoprotein B in bovine
            herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1)
  JOURNAL   Virology 165 (2), 388-405 (1988)
  MEDLINE   88306231
FEATURES             Location/Qualifiers
     source          1..9756
                     /organism="human herpesvirus 1"
                     /db_xref="taxon:10298"
     gene            6555..9266
                     /gene="gB-1"
     CDS             6555..9266
                     /gene="gB-1"
                     /codon_start=1
                     /product="glycoprotein B"
                     /protein_id="AAA19496.1"
                     /db_xref="GI:330110"
                     /translation="MRQGAARGCRWFVVWALLGLTLGVLVASAAPSSPGTPGVAAATQ
                     AANGGPATPAPPAPGPAPTGDTKPKKNKKPKNPPPPRPAGDNATVAAGHATLREHLRD
                     IKAENTDANFYVCPPPTGATVVQFEQPRRCPTRPEGQNYTEGIAVVFKENIAPYKFKA
                     TMYYKDVTVSQVWFGHRYSQFMGIFEDRAPVPFEEVIDKINAKGVCRSTAKYVRNNLE
                     TTAFHRDDHETDMELKPANAATRTSRGWHTTDLKYNPSRVEAFHRYGTTVNCIVEEVD
                     ARSVYPYDEFVLATGDFVYMSPFYGYREGSHTEHTSYAADRFKQVDGFYARDLTTKAR
                     ATAPTTRNLLTTPKFTVAWDWVPKRPSVCTMTKWQEVDEMLRSEYGGSFRFSSDAIST
                     TFTTNLTEYPLSRVDLGDCIGKDARDAMDRIFARRYNATHIKVGQPQYYLANGGFLIA
                     YQPLLSNTLAELYVREHLREQSRKPPNPTPPPPGASANASVERIKTTSSIEFARLQFT
                     YNHIQRHVNDMLGRVAIAWCELQNHELTLWNEARKLNPNAIASATVGRRVSARMLGDV
                     MAVSTCVPVAADNVIVQNSMRISSRPGACYSRPLVSFRYEDQGPLVEGQVGENNELRL
                     TRDAIEPCTVGHRRYFTFGGGYVYFEEYAYSHQLSRADITTVSTFIDLNITMLEDHEF
                     VPLEVYTRHEIKDSGLLDYTEVQRRNQLHDLRFADIDTVIHADANAAMFAGLGAFFEG
                     MGDLGRAVGKVVMGIVGGVVSAVSGVSSFMSNPFGALAVGLLVLAGLAAAFFAFRYVM
                     RLQSNPMKALYPLTTKELKNPTNPDASGEGEEGGDFDEAKLAEAREMIRYMALVSAME
                     RTEHKAKKKGTSALLSAKVTDMVMRKRRNTNYTQVPNKDGDADEDDL"
BASE COUNT     1662 a   3321 c   3202 g   1571 t
ORIGIN
        1 agtgtacata taccaaccgc ataccagacg cacccgaccc ggagcacctg accgtaagca
       61 tctgtgcctc tcgcagggac cccgcgttgc cggccgccgg ggttcatcgg cgcccccgtg
      121 gttacccggg gggttgtcgg tgaaggggag ggattcattc cccaaccccg gtctccaacc
```

```
 181 ctccccttga ccgtcgccgc cccccccggg attttgacgc tcgggagaca taccttgtcg
 241 ggcgtccgtc gtcgtgccgg gattacctcc gttcgcggac cgattgacaa aaggacatgg
 301 agacaaagcc caagacggca accaccatca aggtccccccc cgggcccctg ggatacgtgt
 361 acgctcgcgc gtgtccgtcc gaaggcatcg agcttctggc gttactgtcg gcacgcagcg
 421 gcgattccga cgtcgccgtg gcgcccctgg tcgtgggcct gaccgtggag agcggctttg
 481 aggccaacgt ggccgtggtc gtgggttctc gcacgacggg gctcggggggt accgcggtgt
 541 ccctgaaact gacgccctcg cactacagct cgtccgtgta cgtctttcac ggcggccggc
 601 acctggaccc cagcacccag gccccgaacc tgacgcgcct ttgcgagcgg gcacgccgcc
 661 attttggctt ttcggactac accccccggc ccggcgacct caaacacgag acgacggggg
 721 aggcgctgtg tgagcgcctc ggcctggacc cggaccgcgc cctcctgtat ctggtcgtta
 781 ccgagggctt caaggaggcc gtgtcgatca acaacacctt tctgcacctg ggaggctcgg
 841 acaaggtaac cataggcggg gcggaggtgc accgcatacc cgtgtacccg ttgcagctgt
 901 tcatgccgga ttttagccgt gtcatcgcag agccgttcaa cgccaaccac cgatcgatcg
 961 gggagaaatt tacctacccg cttccgtttt ttaaccgccc cctcaaccgc ctcctgttcg
1021 aggcggtcgt gggacccgcc gccgtggcac tgcgatgccg aaacgtggac gccgtggccc
1081 gcgcggccgc ccacctggcg tttgacgaaa accacgaggg cgccgccctc cccgccgaca
1141 ttacgttcac ggccttcgaa gccagccagg gtaagacccc gcggggcgag cgcgacggcg
1201 gcggcaaggg cgcggcgggc gggttcgaac agcgcctggc ctccgtcatg gccggagacg
1261 ccgccctggc cctcgagtct atcgtgtcga tggccgtctt tgacgagccg cccaccgaca
1321 tctccgcgtg gccgctgttc gagggccagg acacggccgc ggcccgcgcc aacgccgtcg
1381 gggcgtacct ggcgcgcgcc gcgggactcg tggggggccat ggtatttagc accaactcgg
1441 ccctccatct caccgaggtg gacgacgccg gcccggcgga cccaaaggac cacagcaaac
1501 cctcctttta ccgcttcttc ctcgtgcccg ggacccacgt ggcggccaac ccacaggtgg
1561 accgcgaggg acacgtggtg cccgggttcg agggtcggcc caccgcgccc ctcgtcggcg
1621 gaacccagga atttgccggc gagcacctgg ccatgctgtg tgggttttcc ccggcgctgc
1681 tggccaagat gctgtttttac ctggagcgct gcgacggcgc cgtgatcgtc gggcgccagg
1741 agatggacgt gtttcgatac gtcgcggact ccaaccagac cgacgtgccc tgtaacctat
1801 gcaccttcga cacgcgccac gcctgcgtac acacgacgct catgcgcctc cgggcgcgcc
1861 atccaaagtt cgccagcgcc gcccgcggag ccatcggcgt cttcgggacc atgaacagca
1921 tgtacagcga ctgcgacgtg ctgggaaact acgccgcctt ctcggccctg aagcgcgcgg
1981 acggatccga gaccgcccgg accatcatgc aggagacgta ccgcgcggcg accgagcgcg
2041 tcatggccga actcgagacc ctgcagtacg tggaccaggc ggtccccacg gccatggggc
2101 ggctggagac catcatcacc aaccgcgagg ccctgcatac ggtggtgaac aacgtcaggc
2161 aggtcgtgga ccgcgaggtg gagcagctga tgcgcaacct ggtggagggg aggaacttca
2221 agtttcgcga cggtctgggc gaggccaacc acgccatgtc cctgacgctg gacccgtacg
2281 cgtgcgggcc gtgccccctg cttcagcttc tcgggcggcg atccaacctc gccgtgtacc
2341 aggacctggc cctgagtcag tgccacgggg tgttcgccgg gcagtcggtc gaggggcgca
2401 actttcgcaa tcaattccaa ccggtgctgc ggcggcgcgt gatggacatg tttaacaacg
2461 ggtttctgtc ggccaaaacg ctgacggtcg cgctctcgga ggggcggct atctgcgccc
2521 ccagcctaac ggccggccag acggcccccg ccgagagcag cttcgagggc gacgttgccc
2581 gcgtgaccct ggggtttccc aaggagctgc gcgtcaagag ccgcgtgttg ttcgcggggcg
2641 cgagcgccaa cgcgtccgag gccgccaagg cgcgggtcgc cagcctccag agcgcctacc
2701 agaagcccga caagcgcgtg gacatcctcc tcggaccgct gggctttctg ctgaagcagt
2761 tccacgcggc catcttcccc aacggcaagc ccccgggggtc caaccagccg aacccgcagt
2821 ggttctggac ggccctccaa cgcaaccagc ttcccgcccg gctcctgtcg cgcgaggaca
2881 tcgagaccat cgcgttcatt aaaaagtttt ccctggacta cggcgcgata aactttatta
2941 acctggcccc caacaacgtg agcgagctgg cgatgtacta catggcaaac cagattctgc
3001 ggtactgcga tcactcgaca tacttcatca acacccttac ggccatcatc gcgggtgtccc
3061 gccgtccccc cagcgtgcag gctgcggccg cgtggtccgc gcagggcggg gcgggcctgg
3121 aggccggggc ccgcgcgctg gtggacgccg tggacgcgca tccgggcgcg tggacgtcca
3181 tgttcgccag ctgcaacctg ctgcggcccg tcatggcggc gcgcccccatg gtcgtgttgg
3241 ggttgagcat cagcaagtac tacggcatgg ccggcaacga ccgtgtgttt caggccggga
3301 actgggccag cctgatgggc ggcaaaaacg cgtgcccgct ccttattttt gaccgcaccc
3361 gcaagttcgt cctggcctgt ccccggggccg ggtttgtgtg cgcggcctcg agcctcggcg
3421 gcggagcgca cgaaagctcg ctgtgcgagc agctccgggg cattatctcc gagggcgggg
3481 cggccgtcgc cagtagcgtg ttcgtggcga ccgtgaaaag cctggggccc cgcacccaac
3541 agctgcagat cgaggactgg ctggcgctcc tggaggacga gtacctaagc gaggagatga
3601 tggagctgac cgcacgtgcc ctggagcgcg gcaacggcga gtggtcgacg gacgcggccc
3661 tggaggtggc gcacgaggcc gaggccctag tcagccaact cggcaacgcc ggggaggtgt
3721 ttaactttgg ggattttggc tgcgaggacg acaacgcgac gccgttcggc ggcccggggg
3781 ccccgggacc ggcatttgcc ggccgcaaac gggcgttcca cggggatgac ccgtttgggg
3841 aggggccccc cgacaaaaag ggagacctga cgttggatat gctgtgaggg gttgggggt
3901 gggggaacct agggcggggc ggggaatgtg tgtaaaataa attattgcta cgacatccgt
3961 gcttgtttgt gttccgtgtc tatatctctg ggcgggccgt gattcctctc cgcggtgtct
4021 gggaatagaa cagaaacgca cgcgccgccg actcccggct tgccggtcgg cgggcccgcg
```

```
4081 ggaggccgcc ccgaataggg ggaccccggg gctcagccag acgccgggta gcgtacggac
4141 cgcctcggtc caccgcatac tccggccgcg gtacagatcg gcgccgcgag atggccgccc
4201 cggtgtccga gcccaccgtg gcccgtcaaa agttgttagc cctgctcgga caggtgcaga
4261 cctatgtgtt tcagatagag ctgctccggc ggtgcgaccc ccacatcgga cggggggaagc
4321 tcccccaact gaagctgaac gcgcttcagg tgcgggcgct gcggcgtcgt ctgaggccgg
4381 gcctggaggc ccaggccggg gcctttctca ccccgctgtc ggtcaccctg gagttgctgc
4441 tagagtacgc gtggcgcgag ggcgagcggc tcctgggcag cctggagacg ttcgcgaccg
4501 cgggagacgt cgcggtgttt ttcacggaga ccatgggcct ggcccgaccc tgtccgtatc
4561 accaacgggt caggctggat acgtatggcg ggaccgtcca tatggagctg tgtttcctgc
4621 acgacgtcga gaactttcta aagcagctaa actactgcca cctcatcacc ccctcgcgcg
4681 gcgccaccgc cgcgctggag cgcgttcggg agtttatggt ggggcggtg gggtcgggcc
4741 ttatcgtccc cccggagctt agcgacccgt cccaccccctg cgcggtctgt ttcgaggaac
4801 tgtgtgtgac ggcgaaccag ggggcgacga tcgcccgccg cctggcggac cgtatctgta
4861 accacgtcac ccagcaggcg caggtgcggc tggacgccaa cgagctgcgg cggtacctgc
4921 cccacgccgc cgggctgtcg gacgccgacc gcgcgcgggc gctctccgtg ttggaccatg
4981 cgctggcccg gaccgcgggg ggcgacgggc agccccaccc gtcgcccgag aacgactcgg
5041 tccgcaagga ggccgacgcc ctgctggagg cgcacgacgt gtttcaggcc accacgcccg
5101 gcctgtacgc catcagcgaa ttgcgattct ggctcgcgtc cggcgaccgc gccggccaga
5161 ccaccatgga cgcgtttgcc agcaacctga ccgcgctggc gcggcgcgag ttgcagcagg
5221 agaccgccgc ggtggccgtg gaactggcgc tgttcgggcg gcgggcggag catttcgatc
5281 gcgcgttcgg gagccacctg gcggcgctgg atatggtgga cgccctaata atcggcggtc
5341 aggccacgtc acccgacgat cagatcgagg cgctcatccg cgcgtgctac gaccaccacc
5401 tgacgacgcc gctcttgcgg cgcctcgtca gccccgaaca gtgcgacgag gaggcgctgc
5461 gtcgcgtgct ggcgcgcatg ggggcggggg gcgcggcgga cgcgcccaag ggcggcgcgg
5521 gccccgacga cgacggggac cgtgtcgccg tagaggaagg ggcacggggg ttgggagctc
5581 ccggggggcgg gggcgaggac gaagaccgtc gccgcgggcc cggggggcag gggcccgaga
5641 cgtggggggga catcgccacg caagcggccg cggacgtgcg ggagcgacgg cggctgtacg
5701 cggaccgcct gacgaagcgg tcgttggcca gcctgggggcg ctgcgtccgc gagcagcgcg
5761 gggagctcga gaagatgctg cgggtcagcg tccacggcga ggtgctgccc gcgacgttcg
5821 ccgcggtcgc caacggcttt gcggcgcgcg cgcgcttctg cgccctgacg gcgggcgcgg
5881 gcacggtcat cgacaaccgc tcggcgccgg gcgtgttcga cgcgcaccgg ttcatgcgag
5941 cgtctctcct gcgacaccag gtggacccgg ccctgctccc cagcatcacc catcgcttct
6001 tcgagctcgt caacgggccc ctctttgatc actccaccca cagcttcgcc cagccccca
6061 acaccgcgct gtattacagc gtcgagaacg tggggctcct gccgcacctg aaggaggagc
6121 tcgcccggtt catcatgggg gcgggggctc gggtgctgat tgggccgtca gcgaatttca
6181 gaggttttac tgttttgacg gcatttccgg aataacgccc actcagcgcg ccgcctggcg
6241 atatattcgc gagctgatta tcgccaccac actctttgcc tcggtctacc ggtgcgggga
6301 gctcgagttg cgccgcccgg actgcagccg cccgacctcc gaaggtcgtt accgttaccc
6361 gcccggcgta tatctcacgt acgactccga ctgtccgctg gtggccatcg tcgagagcgc
6421 ccccgacggc tgtatcggcc cccggtcggt cgtggtctac gaccgagacg ttttctcgat
6481 cctctactcg gtcctccagc acctcgcccc caggctacct gacggggggc acgacgggcc
6541 cccgtagtcc cgccatgcgc cagggcgccg cgcgggggtg ccggtggttc gtcgtatggg
6601 cgctcttggg gttgacgctg ggggtcctgg tggcgtcggc ggctccgagt tcccccggca
6661 cgcctggggt cgcggccgcg acccaggcgg cgaacggggg acctgccact ccggcgccgc
6721 ccgcccctgg ccccgcccca acgggggaca cgaaaccgaa gaagaacaaa aaaccgaaaa
6781 acccaccgcc gccgcgcccc gccggcgaca acgcgaccgt cgccgcgggc cacgccaccc
6841 tgcgcgagca cctgcgggac atcaaggcgg agaacaccga tgcaaacttt tacgtgtgcc
6901 cacccccac gggcgccacg gtggtgcagt tcgagcagcc gcgccgctgc ccgacccggc
6961 ccgagggtca gaactacacg gagggcatcg cggtggtctt caaggagaac atcgccccgt
7021 acaagttcaa ggccaccatg tactacaaag acgtcaccgt ttcgcaggtg tggttcggcc
7081 accgctactc ccagtttatg gggatctttg aggaccgcgc ccccgtcccc ttcgaggagg
7141 tgatcgacaa gatcaacgcc aagggggtct gtcggtccac ggccaagtac gtgcgcaaca
7201 acctggagac caccgcgttt caccgggacg accacgagac cgacatggag ctgaaaccgg
7261 ccaacgccgc gacccgcacg agccgggggct ggcacaccac cgacctcaag tacaacccct
7321 cgcgggtgga ggcgttccac cggtacggga cgacggtaaa ctgcatcgtc gaggaggtgg
7381 acgcgcgctc ggtgtacccg tacgacgagt ttgtgctggc gactggcgac tttgtgtaca
7441 tgtccccgtt ttacggctac cgggaggggt cgcacaccga acacaccagc tacgccgccg
7501 accgcttcaa gcaggttgac ggcttctacg cgcgcgacct caccaccaag gcccgggcca
7561 cggcgccgac cacccggaac ctgctcacga ccccaagtt caccgtggcc tgggactggg
7621 tgccaaagcg cccgtcggtc tgcaccatga ccaagtggca ggaggtggac gagatgctgc
7681 gctccgagta cggcggctcc ttccgattct cctccgacgc catatccacc accttcacca
7741 ccaacctgac cgagtacccg ctctcgcgcg tggacctggg ggactgcatc ggcaaggacg
7801 cccgcgacgc catggaccgc atcttcgccc gcaggtacaa cgcgacgcac atcaaggtgg
7861 gccagccgca gtactacctg gccaatgggg gctttctgat cgcgtaccag cccccttctca
7921 gcaacacgct cgcggagctg tacgtgcggg aacacctccg agagcagagc cgcaagcccc
```

```
7981 caaacccac gcccccgccg cccggggcca gcgccaacgc gtccgtggag cgcatcaaga
8041 ccacctcctc catcgagttc gcccggctgc agtttacgta caaccacata cagcgccatg
8101 tcaacgatat gttgggccgc gttgccatcg cgtggtgcga gctgcagaat cacgagctga
8161 ccctgtggaa cgaggcccgc aagctgaacc ccaacgccat cgcctcggcc accgtgggcc
8221 ggcgggtgag cgcgcggatg ctcggcgacg tgatggccgt ctccacgtgc gtgccggtcg
8281 ccgcggacaa cgtgatcgtc caaaactcga tgcgcatcag ctcgcggccc ggggcctgct
8341 acagccgccc cctggtcagc tttcggtacg aagaccaggg cccgttggtc gaggggcagg
8401 tgggggagaa caacgagctg cggctgacgc gcgatgcgat cgagccgtgc accgtgggac
8461 accggcgcta cttcaccttc ggtgggggct acgtgtactt cgaggagtac gcgtactccc
8521 accagctgag ccgcgccgac atcaccaccg tcagcacctt catcgacctc aacatcacca
8581 tgctggagga tcacgagttt gtccccctgg aggtgtacac ccgccacgag atcaaggaca
8641 gcggcctgct ggactacacg gaggtccagc gccgcaacca gctgcacgac ctgcgcttcg
8701 ccgacatcga cacggtcatc cacgccgacg ccaacgccgc catgttcgcg ggcctgggcg
8761 cgttcttcga ggggatgggc gacctggggc gcgcggtcgg caaggtggtg atgggcatcg
8821 tgggcggccgt ggtatcggcc gtgtcgggcg tgtcctcctt catgtccaac ccctttgggg
8881 cgctggccgt gggtctgttg gtcctggccg gcctggcggc ggctttcttc gcctttcgct
8941 acgtcatgcg gctgcagagc aaccccatga aggccctgta cccgctaacc accaaggagc
9001 tcaagaaccc caccaacccg gacgcgtccg gggagggcga ggagggcggc gactttgacg
9061 aggccaagct agccgaggcc cgggagatga tacggtacat ggccctggtg tctgccatgg
9121 agcgcacgga acacaaggcc aagaagaagg gcacgagcgc gctgctcagc gccaaggtca
9181 ccgacatggt catgcgcaag cgccgcaaca ccaactacac ccaagttccc aacaaagacg
9241 gtgacgccga cgaggacgac ctgtgacggg gggtttgttg taaataaaaa ccacgggtgt
9301 taaaccgcat gtgcatcttt tggtttgttt gtttggtacg ccttttgtgt gtgtgggaag
9361 aaagaaaagg gaacacataa actcccccgg gtgtccgcgg cctgtttcct ctttcctttc
9421 ccgtgacaaa acggacccccc ttggtcagtg ccgattcccc ccccacgcct tcctccacgt
9481 cgaaggcttt tgcattgtaa agctacccgc ctacccgcgc ctcccaataa aaaaagaaca
9541 tacaccaatg ggtcttattt ggtattacct ggtttattta aaaagatata cagtaagaca
9601 tcccatggta ccaaagaccg gggcgaatca gcgggccccc atcatctgag agacgaacaa
9661 atcggcggcg cgggccgtgt caacgtccac gtgtgctgcg ctgctggcgt tgacaagggc
9721 cccggcctcc gcgttggatg cctccggttg ggatcc
```

//

Revised: October 24, 2001.